



Published in final edited form as:

Magn Reson Med. 2019 October ; 82(4): 1452–1461. doi:10.1002/mrm.27771.

Network Accelerated Motion Estimation and Reduction (NAMER): Convolutional neural network guided retrospective motion correction using a separable motion model

Melissa W. Haskell^{*,1,2}, Stephen F. Cauley^{1,3}, Berkin Bilgic^{1,3}, Julian Hossbach⁴, Daniel N. Splitthoff⁴, Josef Pfeuffer⁴, Kawin Setsompop^{1,3,5}, and Lawrence L. Wald^{1,3,5}

¹A. A. Martinos Center for Biomedical Imaging, Department of Radiology, MGH, Charlestown, MA, United States

²Graduate Program in Biophysics, Harvard University, Cambridge, MA, United States

³Harvard Medical School, Boston, MA, United States

⁴Siemens Healthcare, Erlangen, Germany

⁵Harvard-MIT Division of Health Sciences and Technology, MIT, Cambridge, MA, United States

Abstract

Purpose: We introduce and validate a scalable retrospective motion correction technique for brain imaging that incorporates a machine learning component into a model-based motion minimization.

Methods: A convolutional neural network (CNN) trained to remove motion artifacts from 2D T₂-weighted RARE images is introduced into a model-based data-consistency optimization to jointly search for 2D motion parameters and the uncorrupted image. Our separable motion model allows for efficient intra-shot (line-by-line) motion correction of highly corrupted shots, as opposed to previous methods which do not scale well with this refinement of the motion model. Final image generation incorporates the motion parameters within a model-based image reconstruction. The method is tested in simulations and in vivo motion experiments of in-plane motion corruption.

Results: While the CNN alone provides some motion mitigation (at the expense of introduced blurring), allowing it to guide the iterative joint-optimization both improves the search convergence and renders the joint-optimization separable. This enables rapid mitigation within shots in addition to between shots. For 2D in-plane motion correction experiments, the result is a significant reduction of both image space RMSE in simulations, and a reduction of motion artifacts in the in vivo motion tests.

Conclusion: The separability and convergence improvements afforded by the combined CNN +model-based method show the potential for meaningful post-acquisition motion mitigation in clinical MRI.

*corresponding author. mhaskell@nmr.mgh.harvard.edu.

Keywords

magnetic resonance; motion correction; machine learning; deep learning; convolutional neural networks; image reconstruction

Introduction

Since its inception MRI has been hindered by artifacts due to patient motion, which are both common and costly [1]. Many techniques attempt to track or otherwise account and correct for motion during MRI data acquisition [2], [3], yet few are currently used clinically due to workflow challenges or sequence disturbances. While prospective motion correction methods measure patient motion and update the acquisition coordinates on the fly, they often require sequence modifications (insertion of navigators [4], [5]) or external hardware tracking systems [6], [7]. On the other hand, retrospective methods correct the data after the acquisition, possibly incorporating information from navigators [8] or trackers. Data-driven retrospective approaches operating without tracker or navigator input are attractive because they minimally impact the clinical workflow [9]–[13]. These algorithms estimate the motion parameters providing the best parallel imaging model agreement [14] through the addition of motion operators in the encoding [15]. Unfortunately, this approach typically leads to a poorly conditioned, nonconvex optimization problem. In addition to potentially finding a deficient local minimum, the reconstructions often require prohibitive compute times using standard vendor hardware, limiting the widespread use of the methods.

Machine learning (ML) techniques provide a potential avenue for dramatically reducing the computation time and improving the convergence of retrospective motion correction methods. Recent work has demonstrated how ML can be used to detect, localize, and quantify motion artifacts [16] and deep networks have been trained to reduce motion artifacts [17]–[20]. While the reliance on an ML approach alone shows promise, issues remain with the degree of artifact removal and the robustness of the process to the introduction of blurring or non-physical features. In this work, we attempt to harness the power of a CNN within a controlled model-based reconstruction. We demonstrate how ML can be effectively incorporated into retrospective motion correction approaches based on data consistency error minimization. Other works have balanced data consistency error with ML generated image priors (created using variational networks) to dramatically reduce reconstruction times and improve image quality for highly accelerated acquisitions [21]–[25]. Here, we demonstrate the effective use of ML to guide each step in an iterative model-based retrospective motion correction.

Specifically, we show how a CNN trained to remove motion artifacts from images can improve a model-based motion estimation. During each pass of the iterative process, a CNN image estimate is used as an image prior for the motion parameter search. This motion estimate is then used to create an improved image which can be propagated as input to the CNN to initialize the next pass of the algorithm. The quality of this image estimate can significantly improve the conditioning and convergence of the nonconvex motion parameter optimization. In addition, we demonstrate that with this high quality CNN image estimate,

the motion optimization becomes separable. The separability of the motion model allows for small sub-problems to be optimized in a highly parallel fashion. This allows for a scalable extension of the model-based approach to include intra-shot (line-by-line) motion as well as the inter-shot motion in the RARE acquisition. The increased computation speed also facilitates implementation on standard vendor computation hardware.

Methods

An overview of the Network Accelerated Motion Estimation and Reduction (NAMER) method is shown in Fig. 1. The method is initialized with a SENSE [26] reconstruction of the raw k-space data. The motion correction algorithm is divided into three processes that are iteratively performed: (1) an artifact detecting CNN (Fig. 2) is applied to identify motion artifacts that are subsequently removed, (2) based upon the CNN output, a non-linear optimization estimates the associated motion parameters that minimize the data-consistency error of the forward model, and (3) a model-based reconstruction is performed to generate an updated image based on the motion parameters found in step (2). These steps are then repeated using the updated model-based reconstruction to further reduce the data consistency error and related motion artifacts. Example NAMER code can be found at https://github.com/mwhaskell/namer_mri.

MRI Acquisition

T2-weighted RARE/TSE/FSE [27] data were acquired on 3T MAGNETOM Skyra and Prisma scanners (Siemens Healthcare, Erlangen, Germany) using the product 32-channel head array coil and default clinical protocols. The imaging parameters are: TR/TE=6.1s/103ms, in-plane FOV=220×220mm², 4mm slices, 448×448×35 matrix size, 80% phase resolution, R=2 uniform undersampling, and echo train length (ETL)=11. A 12s FLASH [28] scan provides motion robust auto-calibration data that is used to generate coil sensitivity maps (calculated using ESPRiT [29] from the BART toolbox [30]). In this work data from six healthy subjects were acquired in compliance with institutional practice. Data from four of the subjects were utilized to train the CNN, and the data from the two other subjects were used to evaluate the performance of our method through simulations and supervised motion experiments.

Training Data Generation

Training data for the CNN were created by manipulating raw k-space data (free of motion contamination) using the forward model described in [13] to simulate the effects of realistic patient motion trajectories. Motion trajectories were created using augmentation (shifting and scaling) of timeseries registration information from fMRI scans of patients with Alzheimer's disease (see Supplementary Figure S1). The residual learning CNN attempts to identify the motion artifacts using an L₂-norm loss function against the motion corrupted input image minus the ground truth image. Ten evenly spaced slices from the four healthy subjects were used, and each slice was corrupted by 10 different motion trajectories. Thus, there were 400 motion examples available to choose from in order to create the training data.

The training data was refined through the exclusion of motion corrupted images with RMSE (compared to ground truth) that was greater than 0.50 or less than 0.12. Images with RMSE greater than 0.50 were excluded because they contained severe motion corruption artifacts that could bias the training (due to the large error). Images with RMSE less than 0.12 were excluded because they contained so few motion artifacts which were not productive toward training the CNN to detect artifacts. Overall 76 of the 400 examples were excluded, 61 for $RMSE < 0.12$ and 15 for $RMSE > 0.50$, leaving 324 remaining. To reduce the memory and computational footprint, 24 random cases were dropped to limit the training size to 300 motion cases. These 300 images were divided into patches of size 51×51 with a stride size of 10 (resulting in an overlap region of 41 voxels), which produced 1600 patches per motion corrupted image. For each image, 1250 out of 1600 patches were randomly selected to bring the total number of motion corrupted patches to 375k. From this set of patches, 300k were used for training and 75k used for validation (80/20 training/validation split).

Motion Artifact Detecting Convolutional Neural Network

Fig. 2 shows the 27 layer network topology, which follows previous work [31] and is implemented in Keras [32]. The network isolates the image artifacts within the two-channel (real and imaginary) motion corrupted input patch. The initial layer is a convolutional layer followed by a ReLU activation. The next 25 layers consist of a convolutional layer with batch normalization and ReLU activation, and the final layer is a convolution layer. The number of hidden layers was chosen because the loss function did not improve using more than 25 layers. Additional layers were not added to avoid potential problems with overfitting. Each convolutional layer uses 3×3 kernels with 64 filters. The network was trained using the Adam optimizer [33], with learning rate = $1e-4$ and a mean squared error loss function.

Before being passed through the CNN, the images are scaled to range in magnitude from 0 to 1. Patches are then created, with size 51×51 (as in the training data) and a stride size of 8, which generated 2500 patches per image. After patches were passed through the CNN, they were combined and normalized by the number of patches overlapping at each pixel. No image padding was used for the examples in this paper since the FOV was not tight on the anatomy. The updated artifact free image, x_{cnn} , can be described mathematically as:

$$x_{cnn} = x - CNN(x)$$

where $CNN(x)$ is an image containing the detected motion artifacts within the input image x .

Motion Parameter Optimization

The vector containing the motion parameters, θ , is estimated from x_{cnn} through a non-linear optimization to minimize the data consistency error between the acquired data and the forward model described by the encoding matrix. The encoding includes the effect of the motion trajectory as well as the Fourier encoding and undersampling. The encoding matrix, E_{θ} for the motion parameters θ , is described mathematically as:

$$E_{\theta} = UFCT_{\theta}R_{\theta} \quad (1)$$

where R_{θ} is the rotation operator, T_{θ} is the translation operator, C applies coil sensitivities, F is Fourier encoding, and U is the sampling operator. The motion parameters are found by minimizing the data consistency error between the acquired k-space data, s , and the k-space data generated by applying the motion forward model E_{θ} to the CNN image:

$$[\hat{\theta}] = \operatorname{argmin}_{\theta} \| s - E_{\theta}x_{cnn} \|_2 \quad (2)$$

The minimization is performed using a quasi-Newton search available with the built in `fminunc` function in MATLAB (Mathworks, Natick, MA).

As discussed in prior work [11], [13], the underlying image and motion parameters are tightly coupled which prohibits the separation of the optimization variables into orthogonal subsets. This directly limits the performance of alternating methods to only 1 or 2 productive steps during each alternating pass. However, we will demonstrate that the application of the CNN allows for the efficient decomposition of the optimization. Specifically, the motion parameters, θ , which are typically optimized in a single cost function as shown in (2), can be separated to create a set of much smaller optimization problems, where we can independently estimate θ for each shot of the RARE sequence (11 k-space lines in our case). The motion parameters can then be indexed by the shot number; $\theta = [\theta_1, \theta_2, \dots, \theta_N]$ where N is the total number of shots, and each vector θ_n contains the six rigid body motion parameters for a given shot. The motion forward model, E_{θ} is reduced to only generate the subset of k-space associated with shot n , and is denoted as E_{θ_n} . Similarly, the acquired data for a single shot is denoted as s_n and the cost function for a single shot is:

$$[\hat{\theta}_n] = \operatorname{argmin}_{\theta_n} \| s_n - E_{\theta_n}x_{cnn} \|_2 \quad (3)$$

By using (3) instead of (2), the number of unknowns for any optimization is decreased by a factor of N (the total number of shots). These separate minimizations can be done in parallel, which greatly improves the computational scalability of the retrospective motion correction approach.

This computational efficiency allows us to consider further refinement of the optimization variables. Namely, we extend cost function (3) to consider motion within the RARE shots (intra-shot motion) by assigning additional motion parameters to the individual lines of k-space within a shot. Thus θ_n is expanded from size 6×1 to size $(6 * \text{ETL}) \times 1$, and can be written as $\theta_n = [\theta_{n,1}, \theta_{n,2}, \dots, \theta_{n,L}]$ where L is the number of lines per shot (the ETL of the RARE sequence) and $\theta_{n,l}$ contains the six rigid body motion parameters for shot n , line l . The forward model is further reduced to $E_{\theta_{n,l}}$ to generate the k-space data only for line l of

shot n , and the signal for that line is written as $s_{n,l}$. To find the optimal motion parameters for an individual line, the cost function becomes:

$$[\hat{\theta}_{n,l}] = \operatorname{argmin}_{\theta_{n,l}} \| s_{n,l} - E_{\theta_{n,l}} x_{cnn} \|_2 \quad (4)$$

The ability to selectively refine the optimization variables is extremely valuable as only small portions of the acquisition are likely to suffer from significant intra-shot motion. These shots show significant data consistency error and are often considered as outliers to be discarded [12]. However, with the separable model approach we can expand the model dynamically. The parallel optimizations in Eqn. (3) can be performed for each shot, and then shots that had large data inconsistency after this first pass can then be improved upon using Eqn. (4) (see Supplementary Figure S2). By using this multiscale optimization for the motion parameters, the length of θ can vary depending on how many shots require intra-shot correction, with the minimum length being $6N$ and a maximum length of $6NL$. Similar to Eqn. (2), a quasi-Newton search using MATLAB's `fminunc` is performed for equations (3) and (4).

Model-based Image Reconstruction

Using the raw k-space data, s , and the current estimate for the motion trajectory, $\hat{\theta}$, the linear least squares image reconstruction problem is solved using the conjugate gradient method to find the image, x :

$$[\hat{x}] = \operatorname{argmin}_x \| s - E_{\hat{\theta}} x \|_2 \quad (5)$$

The model-based reconstruction, \hat{x} , can then be fed back through the CNN to identify remaining artifacts. The three steps (apply CNN, motion search, solve for image) are repeated until a stopping criterion is met (either a maximum number of iterations or the change in \hat{x} is below a given threshold). In all cases, the final image returned by NAMER is the model-based reconstruction from Eqn. (5).

Simulation Experiments

The performance of the motion detecting CNN was tested using simulated motion corrupted data. For an unseen subject, each slice was motion corrupted and passed through the CNN. Image space RMSE was calculated for all the motion corrupted slices. Next, a single slice was chosen from the imaging volume and NAMER was performed using both a single cost function (Eqn. (2)), and a separable cost function (Eqn. (3) for all shots). Motion correction using an alternating minimization (only the 2nd and 3rd steps of the NAMER loop) was also performed with both a single and separable cost function, in order to compare to previous methods [11]. No intra-shot motion corruption or correction was used for the comparison experiment. We performed 20 total iterations for all four methods, and the algorithm convergence and final images are compared. Additionally, a simulation using NAMER with the separable cost function was performed in a case with intra-shot motion corrupted data.

We also investigate the potential of combining NAMER with previously developed reduced modeling strategies [13] using simulations. During the third step of NAMER (update of the voxel values), we only updated a strip of voxels that corresponded to the width of a single patch (51 voxels), and then passed the partially updated image through the CNN at the next step. We then evaluated the total number of voxel updates required using a full volume solve versus a targeted voxel update during step 3 of NAMER.

In Vivo Experiments

NAMER was applied to a single brain volume from a healthy subject who was instructed to move during the scan. The subject was asked to shake their head in a “no” pattern for a few seconds, approximately halfway through the 2-minute scan. This motion pattern was used in order to restrict motion artifacts to those caused by within slice motion. Here, a 2D implementation of NAMER, correcting for in-plane translations and rotations, was employed. First, eight iterations of NAMER were performed for each slice, assuming motion only occurred between the shots. Next, one iteration of line-by-line motion correction was performed for the three shots that occurred during the middle of the scan (shots 10, 11, and 12 out of 17 total) and had larger data consistency error.

Results

We present results demonstrating the motion mitigation capabilities of NAMER in simulations and a supervised motion experiment. First, we show that the CNN removes motion artifacts and decreases image space RMSE in all slices of an unseen subject with simulated motion corruption. Next, we show in simulations that NAMER removes more motion artifacts than an alternating minimization for both a single cost function (Eqn. 2) and a separable cost function (Eqn. 3). Unlike the alternating method, the performance of NAMER does not suffer when using a separable cost function. Finally, NAMER removes inter-shot and intra-shot motion artifacts from an in vivo motion experiment where the subject was instructed to move during the scan.

Performance of convolutional neural network

Fig. 3 shows the artifact removal performance of the CNN when applied to simulated motion corruption of data from a subject not included in the training data. Significant ringing artifacts are removed for all slices. When compared to the ground truth images, the simulated motion resulted in an average RMSE of 20% across the slices. The average error was reduced to 16.1% through the application of the CNN. As can be seen in Fig. 3-C, the network removes a significant portion of the motion artifacts, but some residual ringing and blurring are present.

Simulated motion correction using NAMER

Fig. 4 shows the performance of NAMER motion correction compared to an alternating method. Each approach was applied to simulated motion corruption which produced a 43.8% image space RMSE (compared to the ground truth). The final images returned by NAMER for the single cost function (Eqn. 2) and separable cost function (Eqn. 3) formulations show negligible remaining artifacts and the image space RMSE decreased to

12.7% and 11.7% respectively. For the alternating method, significant artifacts remain after the 20 optimization steps. The lack of separability of the alternating optimization is observed as the single cost function outperformed the separable cost function. In that case, the single cost function produced cleaner images with a lower final RMSE (28.2% compared to 31.2%).

The convergence rates of the NAMER and alternating optimizations can be seen in Fig 4-C. The initial data consistency error of the motion corrupted image was 17.4% (corresponding to 43.8% image space RMSE), compared to a data consistency error of 9.4% when using the ground truth motion. With NAMER motion correction, the data consistency error was reduced to the ground truth level of 9.4% after 9 iterations for both the single cost function (Eqn. 2) and separable cost functions (Eqn. 3). After 20 iterations, the alternating method was only able to achieve data consistency errors of 10.4% for the single cost function, and 11.6% for the separable cost function.

For the simulated motion corruption example shown in Figure 4, using a reduced model [13] for image reconstruction (instead of a full volume solve at each NAMER iteration) requires 33% fewer voxels updates. Additionally, NAMER simulations of intra-shot motion correction also produced an artifact mitigated image (Supplementary Figure S3).

NAMER motion correction applied to supervised in vivo motion experiment

NAMER motion correction was applied across the full 3D brain volume to correct artifacts produced by the subject shaking their head “no” during the acquisition. Slice 14 from the 35 slice stack is shown in Fig. 5. The image without any motion correction contains substantial ringing artifacts. Using only between shot NAMER correction with a separable cost function (Eqn. 3) resulted in residual ringing. By refining the model to optimize individual lines within the shots that contain the largest data consistency error (Eqn. 4), the ringing artifacts are significantly reduced. The data consistency RMSE (shown in bottom right of the images in Fig. 5) decreased for all slices within the volume. For slices within the brain region there was an average reduction of 0.2% between the no correction reconstruction and the within shot motion correction. Similar qualitative improvements in image quality were observed for the slices not shown in Fig. 5, and two additional slices are presented in Supplementary Figure S4.

For a single slice of real motion corrupted data (subject 2 in Supplementary Figure S2), the first step of NAMER (CNN evaluation) took on average (averaged across the iterations) 10s on a 12GB NVIDIA Tesla P100. The I/O time to interface the CNN to MATLAB required an additional 40s. The second step of NAMER (motion optimization) took on average 2.7 minutes, and the final step of NAMER (full image solve with motion parameters included) took on average 3.6 minutes.

Discussion

In this work we introduce a scalable retrospective motion correction method that effectively integrates a motion artifact detecting CNN within a model-based motion estimation framework. The image estimates provided by the CNN allow for separation of the motion

parameter search into either individual shots (containing 11 k-space lines in our case) or individual k-space lines. The small optimization problems can be efficiently computed in an embarrassingly parallel fashion. This results in a highly scalable algorithm that has the potential for clinical acceptance. In addition, the separability of the motion optimization facilitates efficient refinement of the model to consider motion disturbances that can occur within a shot. These specific shots can be clearly identified due to their large data consistency error and line-by-line motion correction can be applied. The separability afforded by our method allows us to focus on further improving these limited number of troubled regions without incurring a substantial computational burden. The benefits of this model refinement can be clearly observed (e.g. Fig. 5).

The model-based reconstruction presented here also relaxes concerns about the accuracy, robustness, and predictability of the CNN. As can be seen in Fig. 3, the CNN is able to reduce motion artifacts across all slices from a previously unobserved data set. But, following our expectations, the CNN is not able to completely remove all of the artifacts (see Fig. 3, slice 20) and it can introduce undesirable blurring to the images (see Fig. 3, slice 14). However, the CNN output is accurate enough to both improve the convergence of the reconstruction and promote the separability of the motion parameter optimization. Through the inclusion of more diverse and larger training data sets we expect these benefits to grow. In addition, the NAMER method presented in this work utilizes a standard convolution network. Further improvement might be achieved with more sophisticated networks or loss functions [34]. The topology used here could benefit from further optimization of design parameters, where a sensitivity analysis across network attributes could be performed (i.e. patch size, number of hidden layers). It is important to note that the optimal network may not necessarily be the one which has the lowest validation loss during network training, but instead will be the network that best aids in advancing the motion optimization (step 2 of NAMER).

The NAMER motion mitigation method was assessed through both simulations and supervised in vivo motion experiments. These include testing the CNN artifact removal capabilities on simulated 2D motion corrupted data, simulations to show convergence of the method toward a ground truth, and the supervised head shaking experiment where 2D motion occurred parallel to the imaging planes. The reconstruction framework presented here does generalize to 3D motion trajectories and we think similar benefits can be achieved with adequate training of the CNN, although these remain to be demonstrated. For 3D motion correction with NAMER, the 2D CNN developed here can still be used to mitigate some motion artifacts. However, we think training data that incorporates through plane motion effects will increase the performance of the CNN in general patient motion situations. Due to the limited range of rigid body brain motion, we anticipate that a CNN could be trained on small stacks of slices, where the objective is to remove artifacts from the interior voxels in the stack. The challenges we anticipate in extending NAMER to 3D are present in both the CNN step and the model-based motion optimization. First, the performance of the CNN on data corrupted by through-plane motion that includes spin history effects (which are currently not in the simulated training data model) would need to be evaluated. Second, previous model-based motion optimizations [12], [13] did not include

spin history effects into their motion models, and this could hinder NAMER's performance even if the CNN was able to filter out many of these effects.

The NAMER framework can also accommodate state-of-the-art model reduction and data re-weighting strategies. Prior works have also shown that outlier rejection or soft gating can be beneficial for the mitigation of through plane motion artifacts [12], [35], and we see those prior works as complementary to the NAMER method. The outlier rejection strategy from those works could be used to determine which shots need to have intra-shot motion correction (shown in Fig 5). Further, the data interpolation step in [12] could instead be replaced with our motion mitigating CNN, similar to the cascading network approach in [36]. This could potentially relax the requirement of $2\times$ oversampling used in [12] to achieve robust through plane motion correction. NAMER could also be combined with prior work shown in [13] that performed reduced modelling in image space, and preliminary simulation results presented here demonstrate the potential of the approach, with 33% fewer voxel updates required. Replacing the third step of the NAMER algorithm (see Fig 1) with a reduced model image reconstruction could improve algorithm efficiency and in this case the CNN would only need to evaluate the patches that contain updated voxel values.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to acknowledge a GPU grant from NVIDIA, and support from the National Institute of Mental Health, the National Institute of Biomedical Imaging and Bioengineering, and the NIH Blueprint for Neuroscience Research of the National Institutes of Health under award numbers and NIH grants U01MH093765, R01EB017337, P41EB015896, T90DA022759/R90DA023427, and by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1144152. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation. The authors would like to thank David Salat and Jean-Philippe Coutu for the Alzheimer's disease patient fMRI motion.

References

- [1]. Andre JB, Bresnahan BW, Mossa-Basha M, Hoff MN, Patrick Smith C, Anzai Y, and Cohen WA, "Toward quantifying the prevalence, severity, and cost associated with patient motion during clinical MR examinations," *J. Am. Coll. Radiol.*, vol. 12, no. 7, pp. 689–695, 2015. [PubMed: 25963225]
- [2]. Zaitsev M, Maclaren J, and Herbst M, "Motion artifacts in MRI: A complex problem with many partial solutions," *J. Magn. Reson. Imaging*, vol. 42, no. 4, pp. 887–901, 2015. [PubMed: 25630632]
- [3]. Godenschweger F, Kägebein U, Stucht D, Yarach U, Sciarra A, Yakupov R, Lüsebrink F, Schulze P, and Speck O, "Motion correction in MRI of the brain.," *Phys. Med. Biol.*, vol. 61, no. 5, pp. R32–R56, 2016. [PubMed: 26864183]
- [4]. Tisdall MD, Hess AT, Reuter M, Meintjes EM, Fischl B, and Van Der Kouwe AJW, "Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI," *Magn. Reson. Med.*, vol. 68, pp. 389–399, 2012. [PubMed: 22213578]
- [5]. White N, Roddey C, Shankaranarayanan A, Han E, Rettmann D, Santos J, Kuperman J, and Dale A, "PROMO: Real-time prospective motion correction in MRI using image-based tracking.," *Magn. Reson. Med.*, vol. 63, no. 1, pp. 91–105, Jan. 2010. [PubMed: 20027635]

- [6]. Maclaren J, Armstrong BSR, Barrows RT, Danishad KA, Ernst T, Foster CL, Gumus K, Herbst M, Kadashevich IY, Kusik TP, Li Q, Lovell-Smith C, Prieto T, Schulze P, Speck O, Stucht D, and Zaitsev M, "Measurement and Correction of Microscopic Head Motion during Magnetic Resonance Imaging of the Brain," *PLoS One*, vol. 7, no. 11, pp. 3–11, 2012.
- [7]. Ooi MB, Krueger S, Thomas WJ, Swaminathan SV, and Brown TR, "Prospective real-time correction for arbitrary head motion using active markers," *Magn. Reson. Med*, vol. 62, no. 4, pp. 943–954, Oct. 2009. [PubMed: 19488989]
- [8]. Gallichan D, Marques JP, and Gruetter R, "Retrospective correction of involuntary microscopic head movement using highly accelerated fat image navigators (3D FatNavs) at 7T," *Magn. Reson. Med*, vol. 1039, pp. 1030–1039, 2015.
- [9]. Odille F, Vuissoz P-A, Marie P-Y, and Felblinger J, "Generalized Reconstruction by Inversion of Coupled Systems (GRICS) applied to free-breathing MRI," *Magn. Reson. Med*, vol. 60, no. 1, pp. 146–157, Jul. 2008. [PubMed: 18581355]
- [10]. Loktyushin A, Nickisch H, Pohmann R, and Schölkopf B, "Blind retrospective motion correction of MR images.," *Magn. Reson. Med*, vol. 70, no. 6, pp. 1608–18, Dec. 2013. [PubMed: 23401078]
- [11]. Cordero-Grande L, Teixeira RPAG, Hughes EJ, Hutter J, Price AN, and Hajnal JV, "Sensitivity Encoding for Aligned Multishot Magnetic Resonance Reconstruction," *IEEE Trans. Comput. Imaging*, vol. 2, no. 3, pp. 266–280, Sep. 2016.
- [12]. Cordero-Grande L, Hughes EJ, Hutter J, Price AN, and Hajnal JV, "Three-dimensional motion corrected sensitivity encoding reconstruction for multi-shot multi-slice MRI: Application to neonatal brain imaging," *Magn. Reson. Med*, vol. 79, no. 3, pp. 1365–1376, Mar. 2018. [PubMed: 28626962]
- [13]. Haskell MW, Cauley SF, and Wald LL, "TArgeted Motion Estimation and Reduction (TAMER): Data Consistency Based Motion Mitigation for MRI Using a Reduced Model Joint Optimization," *IEEE Trans. Med. Imaging*, vol. 37, no. 5, pp. 1253–1265, 5 2018. [PubMed: 29727288]
- [14]. Fessler J, "Model-Based Image Reconstruction for MRI," *IEEE Signal Process. Mag*, vol. 27, no. 4, pp. 81–89, Jul. 2010. [PubMed: 21135916]
- [15]. Batchelor PG, Atkinson D, Irarrazaval P, Hill DLG, Hajnal J, and Larkman D, "Matrix description of general motion correction applied to multishot images," *Magn. Reson. Med*, vol. 54, no. 5, pp. 1273–1280, 2005. [PubMed: 16155887]
- [16]. Küstner T, Liebgott A, Mauch L, Martirosian P, Bamberg F, Nikolaou K, Yang B, Schick F, and Gatidis S, "Automated reference-free detection of motion artifacts in magnetic resonance images," *Magn. Reson. Mater. Physics, Biol. Med*, vol. 31, no. 2, pp. 243–256, Apr. 2018.
- [17]. Ko J, Lee J, Yoon J, Lee D, Jung W, and Lee J, "Potentials of Retrospective Motion Correction Using Deep Learning: Simulation Results for a Single Step Translational Motion Correction," *ISMRM Work. Mach. Learn Pacific Grove, CA*, p. 50, 2018.
- [18]. Johnson PM and Drangova M, "Motion correction in MRI using deep learning," *ISMRM Work. Mach. Learn Pacific Grove, CA*, p. 48, 2018.
- [19]. Sommer K, Brosch T, Wiemker R, Harder T, Saalbach A, Hall CS, and Andre JB, "Correction of motion artifacts using a multi-resolution fully convolutional neural network," *Proc. 26th Annu. Meet. ISMRM, Paris*, p. 1175, 2018.
- [20]. Pawar K, Chen Z, Shah NJ, and Egan GF, "Motion Correction in MRI using Deep Convolutional Neural Network," *Proc. 26th Annu. Meet. ISMRM, Paris*, p. 1174, 2018.
- [21]. Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, and Knoll F, "Learning a variational network for reconstruction of accelerated MRI data," *Magn. Reson. Med*, vol. 79, no. 6, pp. 3055–3071, Jun. 2018. [PubMed: 29115689]
- [22]. Aggarwal HK, Mani MP, and Jacob M, "MoDL: Model Based Deep Learning Architecture for Inverse Problems," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2018. [PubMed: 28945591]
- [23]. Adler J and Oktem O, "Learned Primal-Dual Reconstruction," *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1322–1332, Jun. 2018. [PubMed: 29870362]

- [24]. Ward H. a, Riederer SJ, Grimm RC, Ehman RL, Felmlee JP, and Jack CR, "Prospective multiaxial motion correction for fMRI," *Magn. Reson. Med*, vol. 43, pp. 459–469, 2000. [PubMed: 10725890]
- [25]. Wang G, Ye JC, Mueller K, and Fessler JA, "Image Reconstruction Is a New Frontier of Machine Learning," *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1289–1296, 2018. [PubMed: 29870359]
- [26]. Pruessmann KP, Weiger M, Scheidegger MB, and Boesiger P, "SENSE: sensitivity encoding for fast MRI," *Magn. Reson. Med*, vol. 42, no. 5, pp. 952–62, Nov. 1999. [PubMed: 10542355]
- [27]. Hennig J, Nauerth A, and Friedburg H, "RARE imaging: a fast imaging method for clinical MR.," *Magn. Reson. Med*, vol. 3, no. 6, pp. 823–833, 1986. [PubMed: 3821461]
- [28]. Haase A, Frahm J, Matthaei D, Hanicke W, and Merboldt KD, "FLASH imaging. Rapid NMR imaging using low flip-angle pulses," *J. Magn. Reson*, vol. 67, no. 2, pp. 258–266, 1986.
- [29]. Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM, Vasanawala SS, and Lustig M, "ESPIRiT - An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magn. Reson. Med*, vol. 71, no. 3, pp. 990–1001, 2014. [PubMed: 23649942]
- [30]. Uecker M, Ong F, Tamir JI, Bahri D, Virtue P, Cheng JY, Zhang T, and Lustig M, "Berkeley Advanced Reconstruction Toolbox," in *Proceedings of the 23rd Annual Meeting of ISMRM Annual Meeting of ISMRM*, 2015, p. 2486.
- [31]. Bilgic B, Cauley SF, Chatnuntawech I, Manhard MK, Wang F, Haskell M, Liao C, Wald LL, and Setsompop K, "Combining MR Physics and Machine Learning to Tackle Intractable Problems," in *Proceedings of the 26th Annual Meeting of ISMRM, Paris*, 2018, p. 3374.
- [32]. Chollet F, "Keras: The python deep learning library," *Astrophysics Source Code Library*, 2018. .
- [33]. Kingma DP and Ba J, "Adam: A Method for Stochastic Optimization," Dec. 2014.
- [34]. Zhao H, Gallo O, Frosio I, and Kautz J, "Loss Functions for Image Restoration With Neural Networks," *IEEE Trans. Comput. Imaging*, vol. 3, no. 1, pp. 47–57, Mar. 2017.
- [35]. Cheng JY, Zhang T, Ruangwattanapaisarn N, Alley MT, Uecker M, Pauly JM, Lustig M, and Vasanawala SS, "Free-breathing pediatric MRI with nonrigid motion correction and acceleration," *J. Magn. Reson. Imaging*, vol. 42, no. 2, pp. 407–420, 2015. [PubMed: 25329325]
- [36]. Schlemper J, Caballero J, Hajnal JV, Price AN, and Rueckert D, "A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction," *IEEE Trans. Med. Imaging*, vol. 37, no. 2, pp. 491–503, 2018. [PubMed: 29035212]

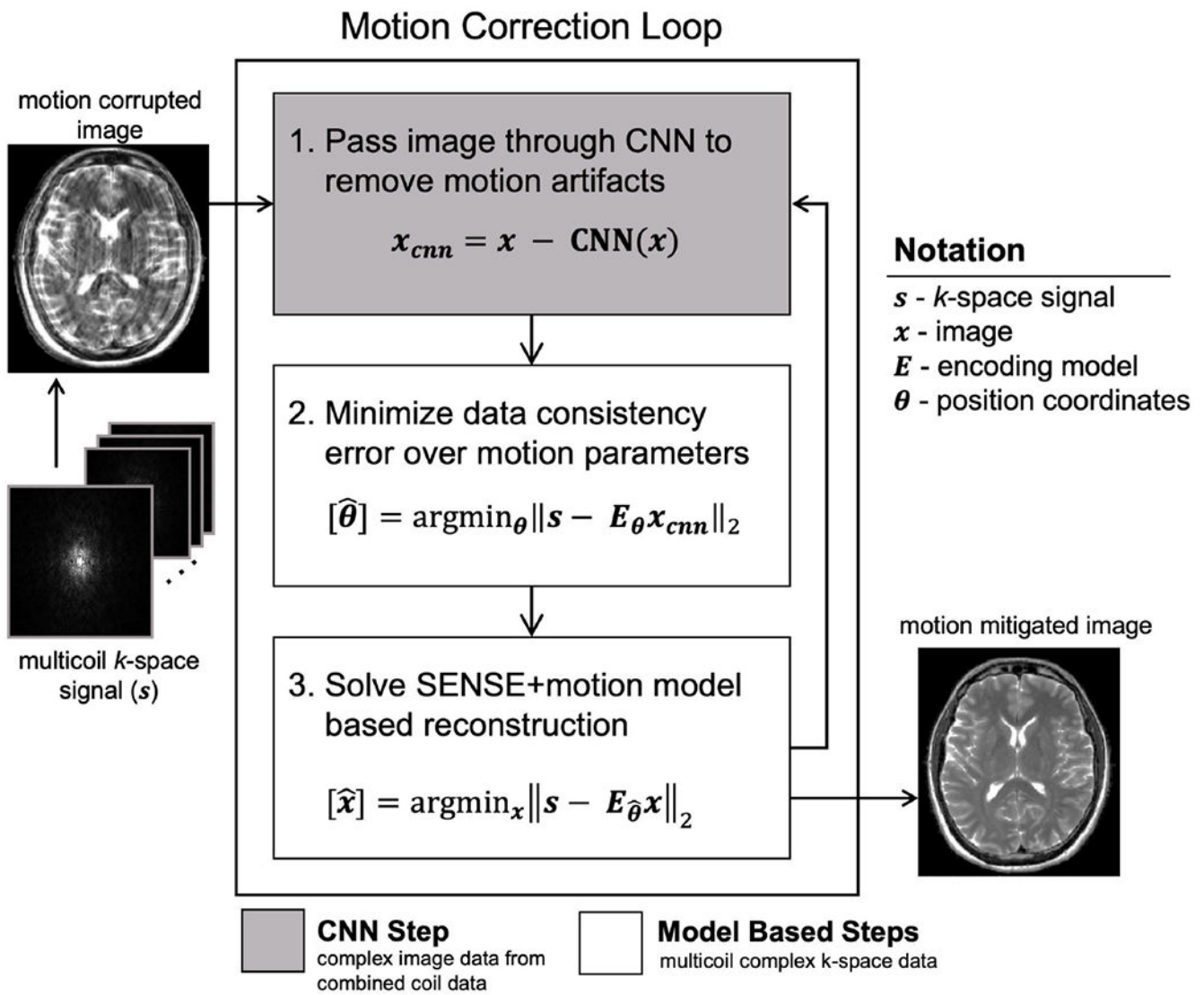


Figure 1. NAMER method overview.

First, a motion corrupted image is reconstructed from the multicoil data, assuming no motion occurred. Next, motion mitigation is performed by looping through three steps: (1) remove motion artifacts in image space by passing the 2-channel complex image (1 channel for the real component and one channel for the imaginary component) through the motion detecting CNN, (2) search for the motion parameters by minimizing the multicoil data consistency error of a motion-inclusive forward model, using complex voxel values from the CNN image, and (3) reconstruct the full image volume using the motion-inclusive multicoil forward model and position coordinates from step (2).

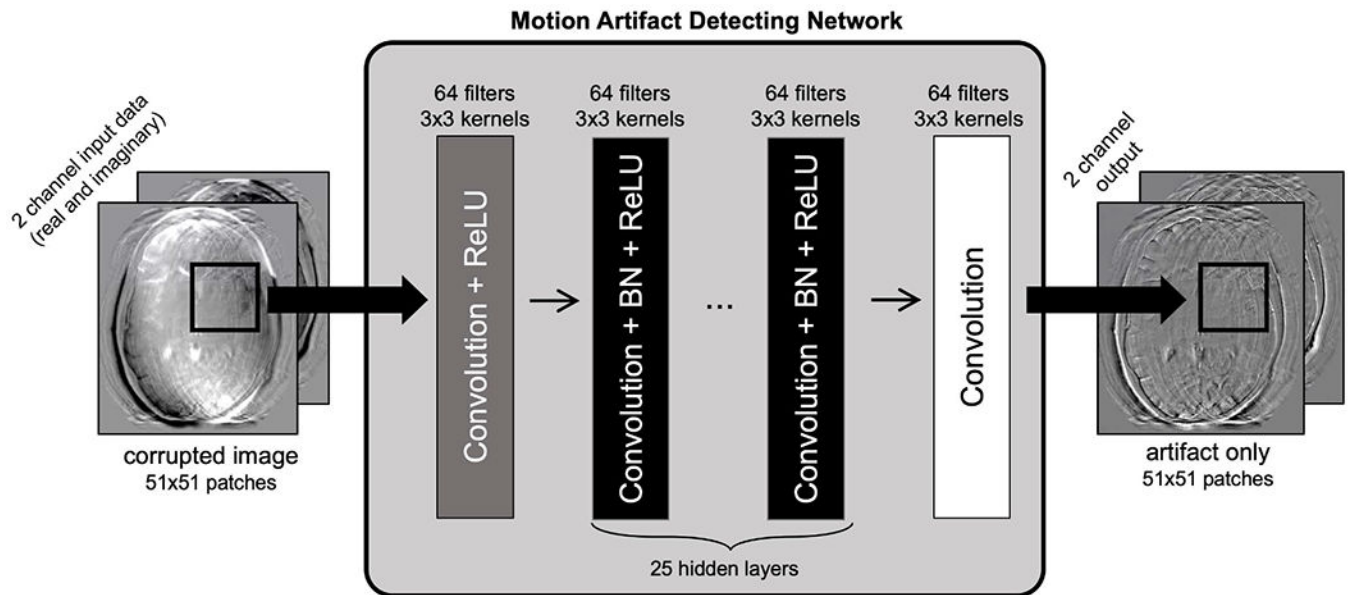


Figure 2. Convolutional neural network for motion artifact detection.

A motion corrupted image is input as two channels (corresponding to the real and imaginary components) to a 27-layer patch-based CNN consisting of convolutional layers, batch normalization, and ReLU nonlinearities. The network outputs the image artifacts, which can be subtracted from the input image to arrive at a motion mitigated image.

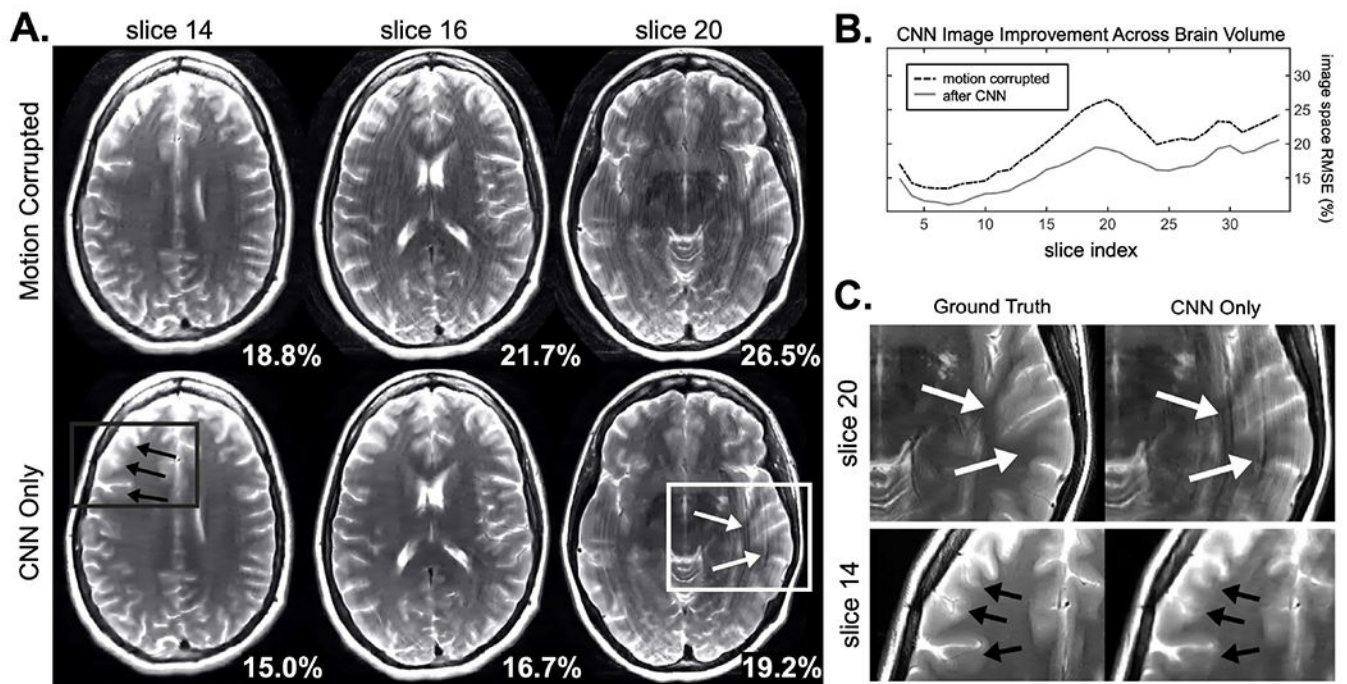


Figure 3. CNN artifact mitigation across whole brain.

A. Representative slices of CNN motion artifact mitigation across the brain volume for a simulated motion example. Bottom right shows the image space RMSE compared to the ground truth image. B. For all slices in the brain volume, the image space RMSE decreased after the CNN, with an average improvement of 3.9%. C. Despite large reductions in artifacts shown in A., compared to the ground truth ringing and blurring artifacts still remain.

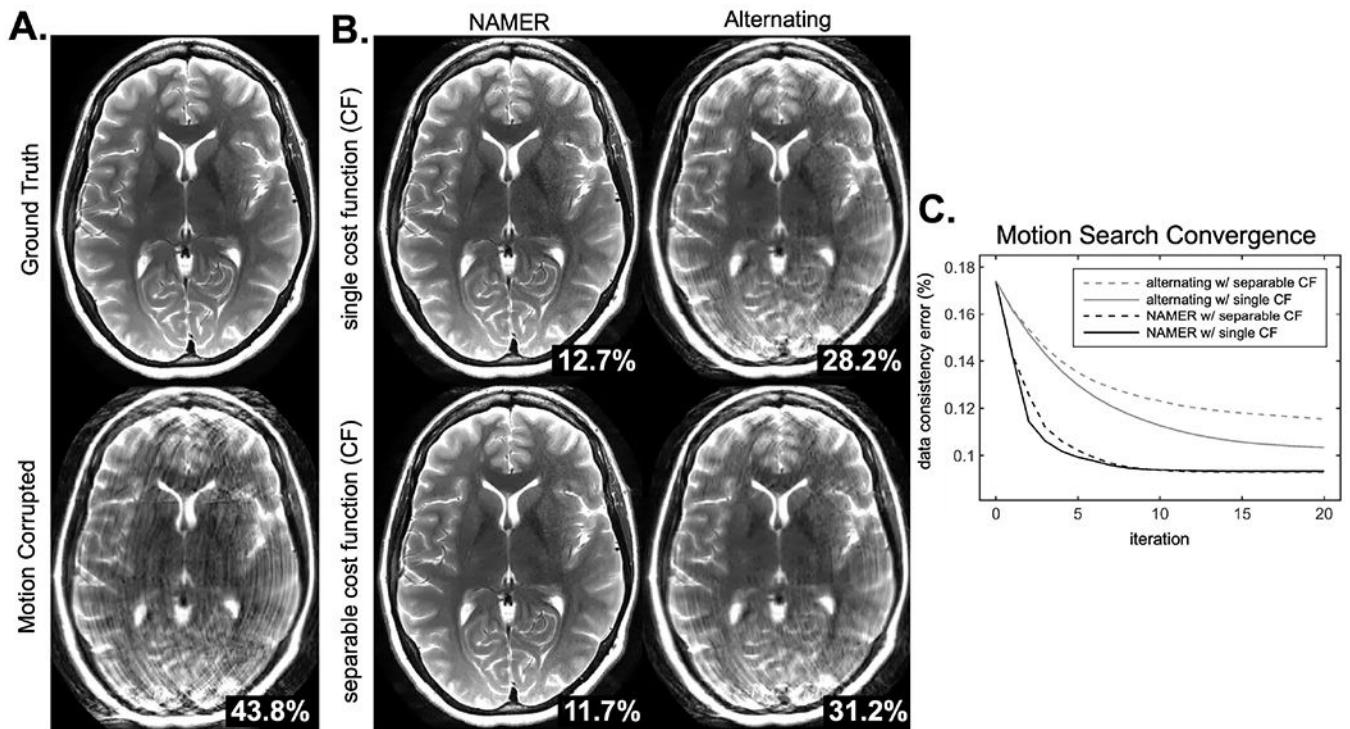


Figure 4. NAMER compared to alternating method in simulated motion data.

A. Original ground truth image and simulated motion corrupted image. B. Image results after 20 iterations of NAMER or an alternating motion correction method. The white numbers in the bottom right are image space RMSE compared to the ground truth image. Top row shows the reconstruction results when using a single cost function for the motion minimization (Eqn. 2), bottom row shows results from a separable cost function for the motion minimization (Eqn. 3). C. Convergence of the four methods displayed in B are shown. Both of the NAMER implementations converge more quickly, and to a lower final data consistency error.

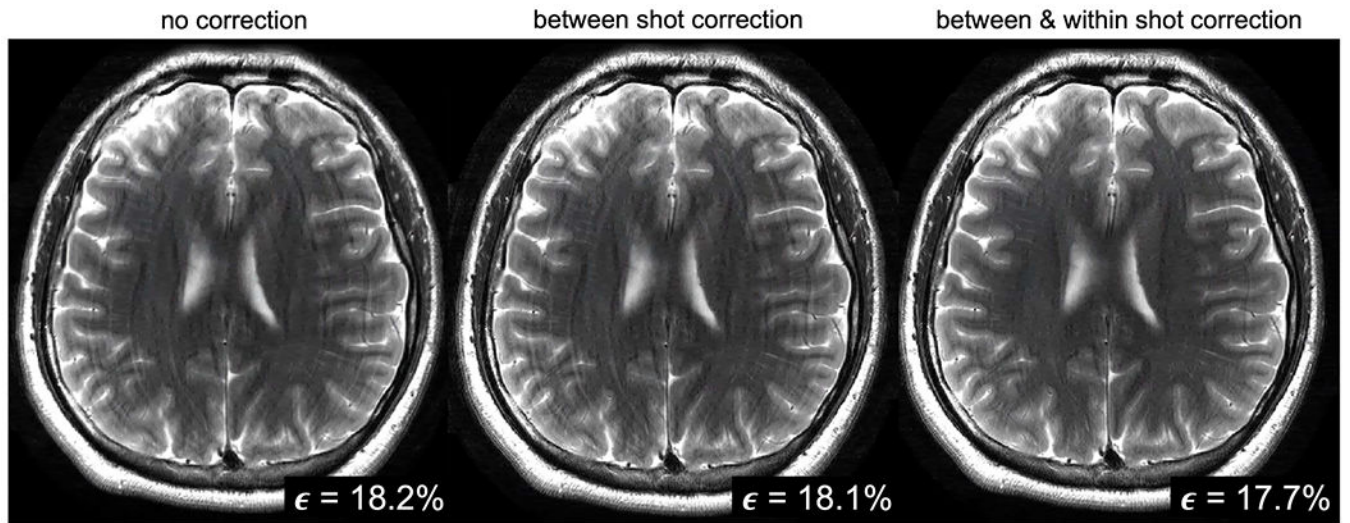


Figure 5. NAMER in vivo inter- and intra-shot motion correction.

Reconstructed images assuming no motion occurred, correcting for motion between shots (inter-shot correction), and correction results after allowing fine-tuning of the motion parameters for each line of k-space at highly corrupted shots (intra-shot correction). Motion artifacts are significantly reduced by allowing the motion parameters to vary across the lines within a shot, as shown in the right column. Data consistency error values are shown in the bottom right of each image.